

REVIEW

Open Access



Cardiac output monitors in septic shock: do they deliver what matters? A systematic review and meta-analysis

Raquel Lamarche-Fontaneto^{1*}, Lee Oud¹, Kayln D. Howell¹, Simran A. Ganeriwal¹, Gaurav U. Manek², Mary C. Schleicher³, Jihad Mallat⁴, Maurizio Cecconi^{5,6}, Abhijit Duggal¹ and Matthew T. Siuba¹

Abstract

To evaluate the interchangeability of cardiac output (CO) monitoring devices compared to reference methods in adult ICU patients with septic shock, we systematically searched electronic databases through January 2025 for prospective studies comparing CO monitors with pulmonary artery catheter (PAC), transpulmonary thermodilution (TPTD), or echocardiography. Eligible studies included Bland–Altman analysis and, when available, trending assessment via polar or 4-quadrant plots, precision, and time response. Agreement was defined as percentage error (PE) < 30%, and acceptable trending as concordance $\geq$ 90%. Pooled bias, limits of agreement (LoA), and PE were calculated using the Sidik–Jonkman random-effects model. Twenty-six studies were included, yielding 37 unique device-reference datasets and encompassing 1,323 patients. PAC was the most common reference (18 datasets), followed by TPTD (16) and echocardiography (3). The pooled bias was 0.15 L min⁻¹ with LoA of $\pm$ 3.45 L min⁻¹ and pooled PE of 49%. Calibrated pulse contour analysis (PCA) showed the best agreement (PE 25%), whereas uncalibrated PCA, thoracic electrical bioimpedance, and bioreactance demonstrated poor agreement (PE $\geq$ 52%). Heterogeneity for mean bias was high across all subgroups (I^2 > 80%). Of 15 datasets reporting trending, only three achieved concordance $\geq$ 90%. Most CO monitors demonstrate poor agreement with reference methods in septic shock. However, their true clinical utility remains unclear, as usual validation frameworks—centered on Bland–Altman analysis—overlook metrics that matter most to intensivists. Precision, time response, and trending ability are critical for real-time decision-making but were rarely assessed. Future studies must incorporate these parameters to meaningfully evaluate device performance at the bedside. PROSPERO registration: CRD42024509384.

Keywords Cardiac output, Hemodynamic monitoring, Meta-analysis, Pulmonary artery catheter, Septic shock

*Correspondence:

Raquel Lamarche-Fontaneto
lamarcr@ccf.org; raquellamarche@gmail.com

¹Department of Pulmonary and Critical Care Medicine, Integrated Hospital Care Institute, Cleveland Clinic, 9500 Euclid Avenue, Cleveland, OH 44195, USA

²Department of Pulmonary and Critical Care Medicine, MedStar Southern Maryland Hospital Center, Clinton, MD, USA

³Cleveland Clinic Floyd D. Loop Alumni Library, Cleveland, OH, USA

⁴Department of Pulmonary and Critical Care Medicine, Integrated Hospital Care Institute, Cleveland Clinic Abu Dhabi, Abu Dhabi, United Arab Emirates

⁵Department of Anesthesia and Intensive Care, IRCCS Humanitas Research Hospital, Milan, Italy

⁶Humanitas University, Milan, Italy



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.



SHOUT
SHARE **IT**

Introduction

Cardiac output (CO) is a primary determinant of oxygen delivery and a cornerstone of hemodynamic management in critical illness. In septic shock, characterized by circulatory instability and evolving perfusion deficits, real-time CO monitoring enables clinicians to assess the effectiveness of resuscitation efforts and tailor ongoing therapy, including fluids and vasoactive agents [1, 2]. The goal is not to reach a specific CO *per se*, but to dynamically inform clinical decisions while minimizing fluid overload and risks of organ dysfunction [3–5]. Accordingly, current guidelines support dynamic CO assessment, particularly in patients who remain unstable after initial fluid resuscitation [6, 7].

Despite its centrality to bedside decision-making, no single CO measurement technique is universally accepted in septic shock [1, 6]. The physiological gold standard, the direct Fick method, is technically demanding and rarely used in contemporary intensive care units (ICU) [8, 9]. Pulmonary artery catheter (PAC) thermodilution and transpulmonary thermodilution (TPTD) are considered validated reference methods, but both require invasive access and offer intermittent measurements [10–13]. Additionally, when combined with pulse contour analysis (PCA), TPTD systems require frequent recalibration to remain reliable in the setting of changing vascular tone [14, 15]. Transthoracic echocardiography (TTE) is another accepted reference technique that provides structural and functional insights [16]. However, it is time consuming, episodic, highly operator-dependent, and sensitive to image quality [17, 18]. Manual assessments, specially with TTE-derived CO, also suffer from significant inter- and intra-observer variability, further compromising reproducibility in unstable patients [16, 18]. While these methods are generally accurate when applied correctly, accuracy alone is insufficient. Intensivists rarely titrate therapy to a fixed CO target; rather, effective management depends on detecting timely, directional trends in response to interventions [19].

To address limitations of traditional reference methods, various minimally invasive and noninvasive CO monitors have emerged. These devices aim to deliver continuous, real-time hemodynamic data while reducing procedural risk [20]. That said, most have been validated primarily in perioperative populations, raising concerns about their generalizability to the dynamic and complex environment of the ICU [21]. Moreover, validation studies have largely focused on agreement metrics—such as mean bias, limits of agreement (LoA), and percentage error (PE)—which assess accuracy but not practical bedside utility [22].

This systematic review and meta-analysis aims to assess these knowledge gaps. We synthesized prospective studies evaluating CO monitoring devices and techniques in adult ICU patients with septic shock, with the dual

objective of quantifying agreement with accepted reference methods and appraising device performance based on precision, response time, and trending ability—key metrics for informing real-time bedside decisions.

Methodology

This review followed Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [23] and was registered in PROSPERO (CRD42024509384). We included prospective studies of CO monitoring in adult ICU patients (≥ 18 years) with septic shock. Studies were eligible if they compared a CO device or technology to a reference method — PAC intermittent thermodilution (PAC-ITD), continuous cardiac output via PAC (PAC-CCO), TPTD, or TTE —and reported agreement metrics. Mixed-population studies were included if $\geq 70\%$ of participants had septic shock or if subgroup data were separately reported. We excluded retrospective studies, conference abstracts, and non-original articles (e.g., reviews, editorials). Only English- or Spanish-language studies were included due to authors' fluency.

We searched Ovid MEDLINE®, Embase, Cochrane CENTRAL, Web of Science, and BIOSIS through January 24, 2025. The strategy combined indexed terms, keywords, and commercial device names using Boolean operators linking CO monitoring, reference methods, and septic shock. Full search details are available in the Electronic Supplementary Material (ESM-1). All results were imported into Covidence for deduplication and screening. Two reviewers (RL, LO) independently screened titles and abstracts, followed by full-text review. Discrepancies were resolved through discussion with a third reviewer (MTS). Reference lists of included articles were also screened, yielding two additional studies [24, 25].

Data extraction

Study-level data were independently extracted by two authors (selected from LO, SAG, KDH, GUM) using a standardized form. A third reviewer (RL) verified all entries for completeness and accuracy. Extracted variables included study design, tested device, reference method, patient characteristics (number of data pairs, type/dose of vasoactive drugs, presence of mechanical ventilation, illness severity scores), hemodynamic interventions, and outcome metrics (mean bias, standard deviation [SD] of bias, LoA, PE, precision, response time, four-quadrant and polar plots, receiver operating characteristic [ROC] analysis). Only reported data were used; missing values were not imputed. When outcome variables were unreported, they were calculated from available data as described in ESM-2.

Risk of bias assessment

Risk of bias was assessed using a modified Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool adapted for method comparison studies (ESM-3) [26]. Three reviewers (LO, KDH, RL) evaluated each study across four domains: patient selection, index test, reference standard, and timing.

Statistical analysis

We analyzed agreement, precision, and trending performance using standardized frameworks to ensure relevance for critical care practice [27–31].

Primary outcome

Agreement between CO monitors and reference methods was assessed using Bland–Altman summary measures. The primary outcome was the pooled PE. A PE < 30% was considered the threshold for acceptable agreement [29] and was calculated as follows:

$$\text{Pooled PE} = \left(\frac{1.96 \times \text{Pooled SD}_{\text{bias}}}{\text{Pooled Mean CO}_{\text{test+reference}}} \right) \times 100$$

Secondary outcomes

Additional agreement metrics

We also pooled (1) the mean bias, reflecting accuracy, (2) the pooled SD of the bias, and (3) the pooled LoA, which—alongside PE—describe inter-patient variability of accuracy. All pooled estimates were calculated using a random-effects Sidik–Jonkman model. Heterogeneity was quantified using the I^2 statistic and categorized as low (25–50%), moderate (50–75%), or high (> 75%). Sub-group analyses were performed by device type (thoracic electrical bioimpedance [TEB], thoracic electrical bioresistance [TEBr], PCA, and pressure recording analytical method [PRAM]) and reference method (PAC, TPTD, and TTE). Sensitivity analyses excluded studies with mixed ICU populations [24, 25, 32, 33]. Meta-regression explored the impact of study-level characteristics including mean CO, systematic vascular resistance (SVR), publication year, blinding, device and reference type, funding source, and geographic region. Publication bias was assessed using funnel plots and Egger’s test (significance: $p < 0.10$ due to the small sample size of individual studies).

Precision

Precision, defined as the random variability of a measurement method, was calculated using the coefficient of variation or error. When reference’s precision was reported, device’s precision was derived from the PE using established methods [27]. In its absence, PE was

interpreted with caution, as it reflects combined imprecision from both test and reference devices [27].

Trending performance

Trending ability was evaluated using four-quadrant and polar plot analyses [30]. A concordance $\geq 90\%$ was considered clinically acceptable [28]. For polar plots, performance was also judged based on angular bias $< \pm 5^\circ$ and radial LoA within $\pm 30^\circ$ based on PAC-ITD’s precision [34]. We assessed the appropriateness of exclusion zones. These should be defined based on the device’s precision, ROC analysis, or established recommendations. For four-quadrant plots—which use Cartesian radius length ($\sqrt{[\Delta\text{CO}_1]^2 + [\Delta\text{CO}_2]^2}$)—a 15% exclusion zone is recommended to filter noise. For polar plots, which use the average of ΔCO vectors, a scaled 10% exclusion zone for relative changes, or 0.5 L min^{-1} for absolute changes, is recommended to account for the $\sim 1.4\times$ shorter radius [28, 30].

Although not a true trending measure, ROC analyses were included as a clinically intuitive tool to assess a device’s ability to detect binary changes, typically using 10% and 15% thresholds for CO and cardiac index (CI). These analyses may indirectly reflect a device’s time response and precision [22].

All statistical analyses were conducted using R (4.5.0), and statistical significance was defined as a two-sided p value < 0.05 .

Results

Following de-duplication, 742 studies were screened. Twenty-six prospective ICU-based observational studies were included, comprising 37 unique datasets (eight studies evaluated multiple devices or software versions; ESM-1). In total, 1,323 patients with septic shock were assessed. The median number of patients per study was 30 (interquartile range [IQR] 18–48), with a median of 135 data pairs (IQR, 73–267). The PRISMA flow diagram is shown in Fig. 1.

Among 23 datasets reporting source of infection, pulmonary ($n = 467$) and abdominal ($n = 167$) infections were most common [35–46]. Most studies reported illness severity scores: the mean SAPS II was 46 (SD 9) [15, 25, 33, 35, 37, 38, 42, 43, 45, 47] and the median APACHE II was 24 (IQR 21–27) [13, 36, 38, 40, 42, 44, 46, 48–51]. Hemodynamic data were variably reported: the median central venous pressure was 14 mmHg (IQR 11–14), median mean arterial pressure (MAP) was 74 mmHg (IQR 69–77), mean SVR was $817 \text{ dyn}\cdot\text{s}\cdot\text{cm}^{-5}$ (SD 168), and mean reference CO was 7.4 L min^{-1} (SD 1.23).

Devices and technologies evaluated included TEB, TEBr, pulse wave transit time (PWTT), PRAM, PCA, double indicator dilution, esophageal doppler, TPTD, PAC-CCO, and TTE. The most common reference

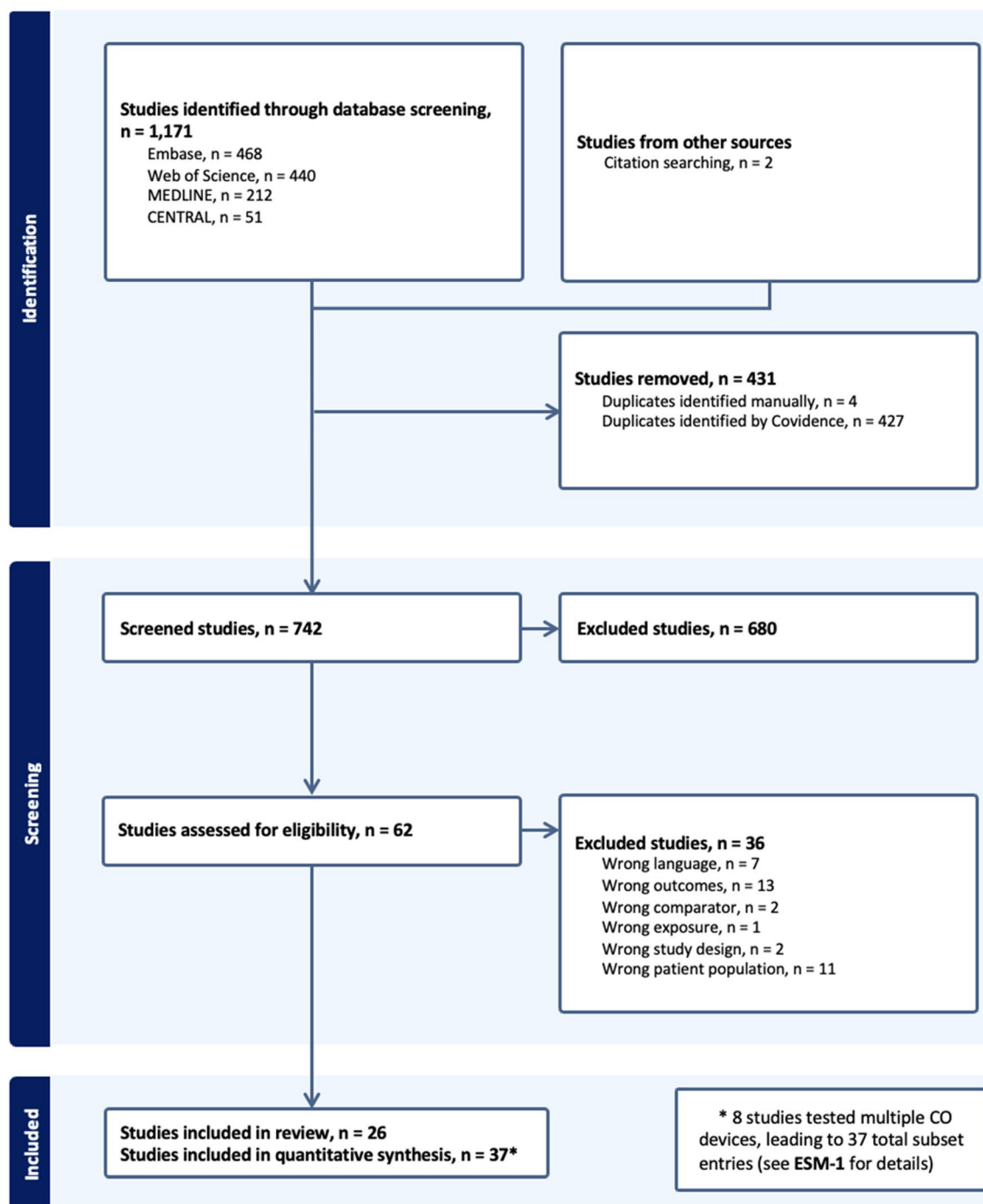


Fig. 1 PRISMA flow diagram of the study selection process

method was PAC (18 datasets: 15 PAC-ITD, 3 PAC-CCO) [36, 37, 39, 41, 44, 46, 47, 49, 52–54], followed by TPTD (16 datasets) [13, 15, 24, 25, 33, 35, 38, 40, 42, 43, 50, 51]. TPTD studies used femoral access, except in two axillary cases from one study [38]. TTE was used as a reference in three studies [32, 45, 48]. At the time of CO measurement, at least 85% of patients were receiving vasoactive support. Study-level characteristics and

trending performance are summarized in Tables 1 and 2. ROC analyses were reported in 14 studies (ESM-5).

Of the 37 datasets, 9 (24%) met the predefined threshold for acceptable agreement ($PE < 30\%$). Trending performance was assessed in 15 datasets, but heterogeneity in ΔCO definitions and exclusion zone thresholds limited the feasibility of meta-analysis. Only three studies

Table 1 Agreement between cardiac output monitoring devices and reference methods in included studies

Study	Device name	Reference method	N/data pairs	Mean CO reference method (SD) [range] (liters min ⁻¹)	Bias (SD) (liters min ⁻¹)	Percentage error (%)	On vasopressors, n (%)	On mechanical ventilation, n (%)
Pressure Recording Analytical Method (PRAM)								
Persona et al. (2021) [35]	MostCare	PiCCO TPTD	40/160	NR	0.11 (0.68)	25	40 (100)	NR
Scoletta et al. (2016)a [32]	MostCare	Echocardiography	47/47	NR	−0.15 (0.63)	24.75	NR	NR
Gopal et al. (2014) [36]	MostCare	PAC-ITD	22/132	6.3 (2.5)	0.31 (1.97)	63	22 (100)	22 (100)
Franchi et al. (2011) [37]	MostCare	PAC-ITD	30/90	7.76 (2.3) [4.1–13.9]	−0.26 (0.98)	24	30 (100)	30 (100)
Thoracic Electrical Bioimpedance (TEB)								
Sharma et al. (2024) [48]	ICON - OSYPKA	Echocardiography	125/375	5.6 (1.8)	2.08 (1.93)	56	125 (100)	125 (100)
Raue et al. (2009) [38]	Aesculon - OSYPKA	PiCCO TPTD	30/30	NR	−0.3 (1.9)	54	27 (90)	29 (97)
Raaijmakers et al. (1998) [39]	Kubicek model	PAC-ITD	13/32	9 (3) [3.3–14.8]	1.8 (2)	49	NR	13 (100)
Raaijmakers et al. (1998) [39]	Sramek–Bernstein model	PAC-ITD	13/32	9 (3) [3.3–14.8]	2.4 (2.8)	69	NR	13 (100)
Thoracic Electrical Bioimpedance (TEBr)								
Galarza et al. (2018)b [33]	Starling-SV	PiCCO TPTD	23/77	6.8	0.06 (1.67)	48 ^e	17 (75)	19 (81)
Kupersztych-Hagege et al. (2013)c [25]	NICOM	PiCCO TPTD	40/144	NR	1.8 (3.26)	82	26 (63)	38 (92)
Pulse Contour Analysis (PCA)								
Khwannimit et al. (2020) [40]	FloTrac/Vigileo 4	PiCCO TPTD	31/156	[2.5–11.3]	0.23 (1.48)	47 ^e	31 (100)	31 (100)
Ganter et al. (2016) [41]	FloTrac/Vigileo 3	PAC-ITD	47/326	6 (2.3) [1.5–19.7]	0.00 (1.09)	36 (95% CI 26–45)	47 (100)	NR
Slagt et al. (2015) [42]	PCA EV1000™ Edwards (8 h post-calibration)	EV1000™ Edwards TPTD	20/267	8.2 (2.3) [2.6–15.9]	0.02 (1.27)	30.3	20 (100)	20 (100)
Slagt et al. (2015) [42]	PCA EV1000™ Edwards (4 h post-calibration)	EV1000™ Edwards TPTD	20/107	8.2 (2.3) [2.6–15.9]	0.05(1.12)	27	20 (100)	20 (100)
Slagt et al. (2015) [42]	FloTrac/Vigileo 3.02	EV1000™ Edwards TPTD	20/301	8.2 (2.3) [2.6–15.9]	−0.86(1.85)	47	20 (100)	20 (100)
Monnet et al. (2015) [43]	FloTrac/Vigileo 3.02	PiCCO TPTD	60/120	NR	1 (1.94)	59	60 (100)	60 (100)
Monnet et al. (2015) [43]	ProAQT Pulsioflex	PiCCO TPTD	60/120	NR	−0.2 (1.43)	39	60 (100)	60 (100)
Monnet et al. (2015) [43]	PCA PiCCO	PiCCO TPTD	60/120	NR	0.22 (0.75)	21	60 (100)	60 (100)
Slagt et al. (2013) [49]	FloTrac/Vigileo 3.02	PAC-ITD	19/314	8.6 (2.7) [3.8–17.3]	1.7 (2.4)	61	19 (100)	19 (100)
Marqué et al. (2013) [47]	FloTrac/Vigileo 3.02	PAC-CCO	18/1,201	6.4 [1.6–13.6]	−0.2 (2.14)	66	18 (100)	NR
Monnet et al. (2012)d [24]	Flotracc/Vigileo 3	PiCCO TPTD	48/120	6.8 (2.6)	0.52 (1.88)	54 ^e	48 (100)	48 (100)
De Backer et al. (2010) [53]	FloTrac/Vigileo 1.14	PAC-ITD	58/401	7.5 (2) [2.7–14.6]	−0.8 (1.12)	31 (95% CI 20–37)	50 (86)	51 (88)
De Backer et al. (2010) [53]	FloTrac/Vigileo 3	PAC-ITD	58/401	7.5 (2) [2.7–14.6]	0 (1.12)	29.7 (95% CI 24–37)	50 (86)	51 (88)
Monnet et al. (2010) [15]	Flotracc/Vigileo 1.10	PiCCO TPTD	80/240	NR	−0.18 (1.88)	61 ^e	53 (66)	NR
Monnet et al. (2010) [15]	PCA PiCCO	PiCCO TPTD	80/240	NR	−0.14 (0.72)	22	53 (66)	NR
Böettger (2010) [50]	FloTrac/Vigileo 1	PiCCO TPTD	19/55	6.35 (1.41)	0.72 (1.47)	48	NR	19 (100)
Slagt et al. (2010) [44]	FloTrac/Vigileo 1.07	PAC-ITD	4/86	[3.6–10.4]	−1.6 (1.6)	48	4 (100)	4 (100)
Slagt et al. (2010) [44]	FloTrac/Vigileo 1.10	PAC-ITD	5/73	[2.9–12.6]	−1.2 (1.1)	32	5 (100)	5 (100)

Table 1 (continued)

Study	Device name	Reference method	N/data pairs	Mean CO reference method (SD) [range] (liters min ⁻¹)	Bias (SD) (liters min ⁻¹)	Percentage error (%)	On vasodilators. n (%)	On mechanical ventilation. n (%)
Spöhr et al. (2007) [52]	PCA PiCCO	PAC-CCO	14/182	8.8 (2.2) [3.6–12.8]	0.1 (1.38)	30	14 (100)	14 (100)
Sakka et al. (2007) [13]	FloTrac/Vigileo 1.07	PiCCO TPTD	24/72	6.6 (1.8) [3.2–10.1]	0.5 (2.3)	73	24 (100)	24 (100)
Jellema et al. (1999) [54]	Modelflow	PAC-ITD	32/189	9.0 (3.0)	−0.1 (0.8)	17.4	32 (100)	32 (100)
Transpulmonary Thermodilution (TPTD)								
Spöhr et al. (2007) [52]	TPTD PiCCO	PAC-CCO	14/182	8.8 (2.2) [3.6–12.8]	0.2 (1.12)	24	14 (100)	14 (100)
Pulse Wave Transit Time (PWTT)								
Feissel et al. (2015) [45]	EsCCO	Echocardiography	25/50	5.27 (1.17)	0.2 (1.12)	37	25 (100)	25 (100)
Double Indicator Dilution								
Raaijmakers et al. (1998) [39]	Iced glucose and indocyanine dye measures CO via transit time to the femoral artery	PAC-ITD	13/32	9 (3) [3.3–14.8]	1.1 (1.3)	27	13 (100)	13 (100)
Continuous Cardiac Output with a Pulmonary Artery Catheter (PAC-CCO)								
Ganter et al. (2016) [41]	PAC-CCO	PAC-ITD	47/326	6 (2.3) [1.5–19.7]	0.23 (1.3)	41 (95% CI 32–50)	47 (100)	NR
De Backer et al. (2010) [53]	PAC-CCO	PAC-ITD	58/401	7.5 (2) [2.7–14.6]	0.7 (1.07)	27 (95% CI 22–34)	50 (86)	51 (88)
Echocardiography								
Spathoulas et al. (2022) [51]	Echocardiography	PiCCO TPTD	14/71	6.64 (NR)	−0.0011 (0.29)	8.6 ^e	14(100)	14(100)
Esophageal Doppler								
Chew et al. (2009) [44]	CardioQ, Deltex Medical	PAC-ITD	12/97	[2.87–11]	0.43 (1.2)	58	12(100)	12(100)

CI confidence interval, NR not reported, PAC-CCO pulmonary artery catheter continuous cardiac output, PAC-ITD pulmonary artery catheter intermittent thermodilution, PE percentage error, TPTD transpulmonary thermodilution

^asubset analysis of sepsis: 47 patients with sepsis, 47 data pairs (N=400)

^b>70% septic study: 23/32 (72%) patients were septic

^c>70% septic study: 40/48 (83%) patients were septic

^d>70% septic study: 48/60 (80%) patients were septic

^ePE calculated as = 1.96 × (standard deviation of bias)/mean CO of reference method

achieved the predefined acceptable concordance threshold ($\geq 90\%$) [37, 40, 43].

Pulse Contour analysis (PCA)

PCA was the most frequently assessed modality (20 datasets), spanning uncalibrated (FloTrac/Vigileo® versions 1–4, ProAQT-Pulsioflex®) and calibrated systems (PCA PiCCO®, PCA EV1000™ Edwards, Modelflow®). Mean bias ranged from −1.6 to 1.7 L min⁻¹. Most studies used synchronized time scales for data collection, though some deviated from this [15, 52]. Four datasets met the predefined agreement threshold (PE < 30%; Table 1). Radial artery access predominated; only two studies assessed access's impact on agreement, finding no significant differences [41, 53].

All calibrated systems met the PE < 30% threshold except one PCA PiCCO® dataset with borderline performance (PE = 30%) using PAC-CCO as reference—a method known to lag by 3–10 min [52]. PCA EV1000™

showed acceptable agreement (PE 21–27%) within 4 to 6 h post-calibration but declined to 30.3% beyond 8 h [42].

All FloTrac/Vigileo® datasets reported PE > 30%. In De Backer et al., PE was 29.7% (95% confidence interval 24–37%) based on offline manufacturer calculations, but poor standalone precision (28–30%) undermined confidence in this estimate [53]. Several studies evaluated agreement under unstable hemodynamic conditions [13, 42, 44, 47, 49], though, Marqué et al. observed improved agreement (PE 36%) when analysis was restricted to stable periods [47].

Precision was reported in three studies. FloTrac/Vigileo® version 3 showed 48% precision compared to 7% for TPTD [42] and versions 2 and 3 showed 28–30% precision versus 12% for PAC-ITD [53]. Version 1.07 exhibited the poorest performance, with ~72% precision compared to 13% for TPTD under dynamic conditions [13].

Table 2 Trending performance of included studies using Polar and Four-Quadrant plot analyses

Study	Device	Reference method	N/data pairs	Intervention	Trending data	On vasoactives <i>n</i> (%)	On mechanical ventilation <i>n</i> (%)
Pressure Recording Analytical Method (PRAM)							
Persona et al. (2021) [35]	MostCare	PiCCO TPTD	40/40	500 mLs crystalloids (<i>n</i> = 15) or vasoactive dose change	Δ CO Polar Plot Exclusion zone: Δ CO < 10% Total data points: 40 Data points after exclusion (<i>n</i>): 11 Concordance: 82% within 30° of polar axis	40 (100)	NR
Franchi et al. (2011) [37]	MostCare	PAC-ITD	30/60	Vasoactive dose change	Δ CO 4-Quadrant Plot Exclusion zone: Δ CO < 0.5 L min ⁻¹ Total data points: 60 Data points after exclusion (<i>n</i>): 39 Concordance: 100%	30 (100)	30 (100)
Thoracic Electrical Bioimpedance (TEBr)							
Galarza et al. (2018)a [33]	Starling-SV	PiCCO TPTD	23/45	Passive leg raise 500 mLs NS over 10 min if preload responsive	% Δ CI 4-Quadrant Plot Exclusion zone: Δ CI < 15% Total data points: 45 Data points after exclusion (<i>n</i>): NR Concordance: 78% Δ CO Polar Plot Exclusion zone: Δ CO < 0.5 L min ⁻¹ Total data points: 45 Data points after exclusion (<i>n</i>): NR Angular bias: 10° Limits of agreement: -23° to 44°	17 (75)	19 (81)
Kupersztynch-Hagege et al. (2013)b [25]	NICOM	PiCCO TPTD	40/144	Passive leg raise 500 mLs of NS over 10 min	% Δ CI 4-Quadrant Plot Exclusion zone: Δ CI < 15% Total data points: 144 Data points after exclusion (<i>n</i>): NR Concordance: 52%	26 (63)	38 (92)
Pulse Contour Analysis (PCA)							
Kh-wann-imit et al. (2020) [40]	FloTrac/Vigileo 4	PiCCO TPTD	31/156	500mLs of NS over 30 min Vasoactive dose change	% Δ CI 4-Quadrant Plot Exclusion zone: Δ CI < 10% Total data points: 156 Data points after exclusion (<i>n</i>): 94 Concordance: 94%	31 (100)	31 (100)
Kh-wann-imit et al. (2020) [40]	FloTrac/Vigileo 4	PiCCO TPTD	30/56	Vasoactive dose change	% Δ CI 4-Quadrant Plot Exclusion zone: Δ CI < 10% Total data points: 56 Data points after exclusion (<i>n</i>): NR Concordance: 96%	30 (100)	30 (100)
Kh-wann-imit et al. (2020) [40]	FloTrac/Vigileo 4	PiCCO TPTD	11/16	500mLs of NS over 30 min	% Δ CI 4-Quadrant Plot Exclusion zone: Δ CI < 10% Total data points: 16 Data points after exclusion (<i>n</i>): NR Concordance: 83%	11 (100)	11 (100)
Ganter et al. (2016) [41]	FloTrac/Vigileo 3	PAC-ITD	47/326	Measurements every 4- hours for 24 h	Δ CO Polar Plot Exclusion zone: Δ CO < 0.5 L min ⁻¹ Total data points: 326 Data points after exclusion (<i>n</i>): 169 Concordance: 85% (poor trending ability) and 70% within 30° of polar axis Angular bias: 3° Limits of agreement: -51° to 58°	47 (100)	NR

Table 2 (continued)

Study	Device	Reference method	N/data pairs	Intervention	Trending data	On vasoactives <i>n</i> (%)	On mechanical ventilation <i>n</i> (%)
Slagt et al. (2015) [42]	PCA EV1000™ Edwards	EV1000™ Edwards TPTD	20/267	"CO measured with changes to therapy with fluids and vaso-active agents"	Δ CO Polar Plot Exclusion zone: Δ CO < 1.2 L min ⁻¹ or < 15% Total data points: 267 Data points after exclusion (n): 39 Concordance: 74% and 68% within 30° of polar axis	20 (100)	20 (100)
Slagt et al. (2015) [42]	FloTrac/Vigileo 3	EV1000™ Edwards TPTD	20/301	"CO measured with changes to therapy with fluids and vaso-active agents"	Δ CO Polar Plot Exclusion zone: Δ CO < 1.2 L min ⁻¹ or < 15% Total data points: 301 Data points after exclusion (n): 40 Concordance: 88% and 76% within 30° of polar axis	20 (100)	20 (100)
Monnet et al. (2015) [43]	FloTrac/Vigileo 3.02 software	PiCCO TPTD	20/40	500 mLs NS over 10 min	% Δ CI Polar Plot Exclusion zone: Δ CI < 15% Total data points: 40 Data points after exclusion (n): NR Concordance: 73%	20 (100)	20 (100)
Monnet et al. (2015) [43]	FloTrac/Vigileo 3.02 software	PiCCO TPTD	40/80	Vasoactive dose change	% Δ CI Polar Plot Exclusion zone: Δ CI < 15% Total data points: 80 Data points after exclusion (n): NR Concordance: 41%	40 (100)	40 (100)
Monnet et al. (2015) [43]	ProAQT Pulsioflex (non-calibrated)	PiCCO TPTD	20/40	500 mLs NS over 10 min	% Δ CI Polar Plot Exclusion zone: Δ CI < 15% Total data points: 40 Data points after exclusion (n): NR Concordance: 91%	20 (100)	20 (100)
Monnet et al. (2015) [43]	ProAQT Pulsioflex (non-calibrated)	PiCCO TPTD	40/80	Vasoactive dose change	% Δ CI Polar Plot Exclusion zone: Δ CI < 15% Total data points: 80 Data points after exclusion (n): NR Concordance: 39%	40 (100)	40 (100)
Monnet et al. (2015) [43]	ProAQT Pulsioflex (calibrated)	PiCCO TPTD	40/80	Vasoactive dose change	% Δ CI Polar Plot Exclusion zone: Δ CI < 15% Total data points: 80 Data points after exclusion (n): NR Concordance: 41%	40 (100)	40 (100)
Monnet et al. (2015) [43]	ProAQT Pulsioflex (calibrated)	PiCCO TPTD	20/40	500 mLs NS over 10 min	% Δ CI Polar Plot Exclusion zone: Δ CI < 15% Total data points: 40 Data points after exclusion (n): NR Concordance: 79%	20 (100)	20 (100)
Monnet et al. (2015) [43]	PCA PiCCO	PiCCO TPTD	40/80	Vasoactive dose change	% Δ CI Polar Plot Exclusion zone: Δ CI < 15% Total data points: 80 Data points after exclusion (n): NR Concordance: 77%	40 (100)	40 (100)
Monnet et al. (2015) [43]	PCA PiCCO	PiCCO TPTD	VE 20/40	500 mLs NS over 10 min	% Δ CI Polar Plot Exclusion zone: Δ CI < 15% Total data points: 40 Data points after exclusion (n): NR Concordance: 85%	20 (100)	20 (100)
Monnet et al. (2012)c [24]	FloTrac/Vigileo 3	PiCCO TPTD	20/40	500 mLs of NS over 30 min	% Δ CI 4-Quadrant Plot Exclusion zone: Δ CI < 12% Total data points: 40 Data points after exclusion (n): NR Concordance: 70%	20 (100)	20 (100)

Table 2 (continued)

Study	Device	Reference method	N/data pairs	Intervention	Trending data	On vasoactives n (%)	On mechanical ventilation n (%)
Monnet et al. (2012)c [24]	FloTrac/Vigileo 3	PiCCO TPTD	40/40	Vasoactive dose change	% Δ CI 4-Quadrant Plot Exclusion zone: Δ CI < 12% Total data points: 40 Data points after exclusion (n): NR Concordance: 63%	40 (100)	40 (100)
Slagt et al. (2013) [49]	FloTrac/Vigileo 3.02	PAC-ITD	19/314	Volume expansion Vasoactive dose change	Δ CO Polar Plot Exclusion zone: Δ CO < 0.5 L min ⁻¹ Total data points: 314 Data points after exclusion (n): 163 Concordance: 85%	19 (100)	19 (100)
Continuous Cardiac Output with a Pulmonary Artery Catheter (PAC-CCO)							
Ganter et al. (2016) [41]	PAC-CCO	PAC-ITD	47/326	Measurements every 4- hours for 24 h	Δ CO Polar Plot Exclusion zone: Δ CO < 0.5 L min ⁻¹ Total data points: 326 Data points after exclusion (n): 155 Concordance: 81% and 71% within 30° of polar axis Angular bias: 0.1° Angular LoA: -57° to 57°.	47 (100)	NR

NR not reported, NS normal saline, PAC CCO pulmonary artery catheter continuous cardiac output, PAC ITD pulmonary artery catheter intermittent thermodilution, TPTD transpulmonary thermodilution, % Δ CO percent changes in cardiac output, % Δ CI percent changes in cardiac index, Δ CO absolute changes in cardiac output, Δ CI absolute changes in cardiac index

^a>70% septic study: 23/32 (72%) patients were septic

^b>70% septic study: 40/48 (83%) patients were septic

^c>70% septic study: 48/60 (80%) patients were septic

Trending was evaluated in 10 datasets (Table 2). Only ProAQ-T-Pulsioflex® [43] and FloTrac/Vigileo® version-4 [40] achieved $\geq 90\%$ concordance in response to vaso-pressor changes. The latter applied a smaller-than-recommended exclusion zone.

Pressure recording analytical method (PRAM)

PRAM, evaluated using the MostCare® device, was assessed in four studies (Table 1). Mean bias ranged from -0.15 to 0.31 L min⁻¹, with all studies applying synchronized time scales during data collection. Only Gopal et al. reported poor agreement (PE 63%) in their pragmatic study conducted during periods of high CO variability, without excluding conditions known to compromise PRAM performance [36].

Trending was assessed in two studies (Table 2) [35, 37]. Franchi et al. reported 100% concordance following a vasopressor-induced MAP increase (70–90 mmHg) using a 0.5 L min⁻¹ exclusion zone in the four-quadrant plot. In contrast, Persona et al. did not meet the $\geq 90\%$ threshold. They applied lower-intensity interventions (norepinephrine changes of 0.03 mcg/kg/min) resulting in few remaining Δ CO data pairs post-exclusion.

Thoracic electrical bioimpedance (TEB) and thoracic electrical bioreactance (TEBr)

TEB and TEBr were assessed in six datasets (Table 1). None met the PE < 30% or trending thresholds. Only one

study reported aligned time scales during data collection [39].

Risk of bias within studies

Quality assessment results are detailed in ESM-4. Only Scolletta et al. was judged at high risk of reference bias due to inclusion of patients, with suboptimal, but adequate TTE windows [28]. The study was retained as all scans were deemed sufficient for analysis [32].

Synthesis

The overall random-effects pooled bias was 0.15 L min⁻¹ with LoA of ± 3.45 L min⁻¹ and a PE of 49.4%. Inter-study heterogeneity was substantial ($I^2 = 97.5\%$, $P < 0.001$) (Fig. 2). Egger's test showed no evidence of publication bias ($P = 0.16$), and the funnel plot appeared symmetrical (ESM-6).

Subgroup analysis revealed high heterogeneity across all device types. For PCA devices, the pooled bias was -0.02 L min⁻¹ with LoA of ± 1.69 L min⁻¹ and a PE of 48.2% ($I^2 = 97\%$). PRAM showed a pooled bias of -0.03 L min⁻¹, LoA of ± 2.78 L min⁻¹, and a PE of 44% ($I^2 = 79\%$). TEB and TEBr demonstrated the poorest agreement, with a pooled bias of 0.71 L min⁻¹, LoA of ± 2.9 L min⁻¹, and a PE of 71.8% ($I^2 = 97.3\%$) (Fig. 3).

Among PCA systems, calibrated PCA showed a pooled bias of -0.02 L min⁻¹, LoA of ± 0.98 L min⁻¹, and a PE of 24.7% ($I^2 = 81.9\%$), while uncalibrated PCA had a pooled

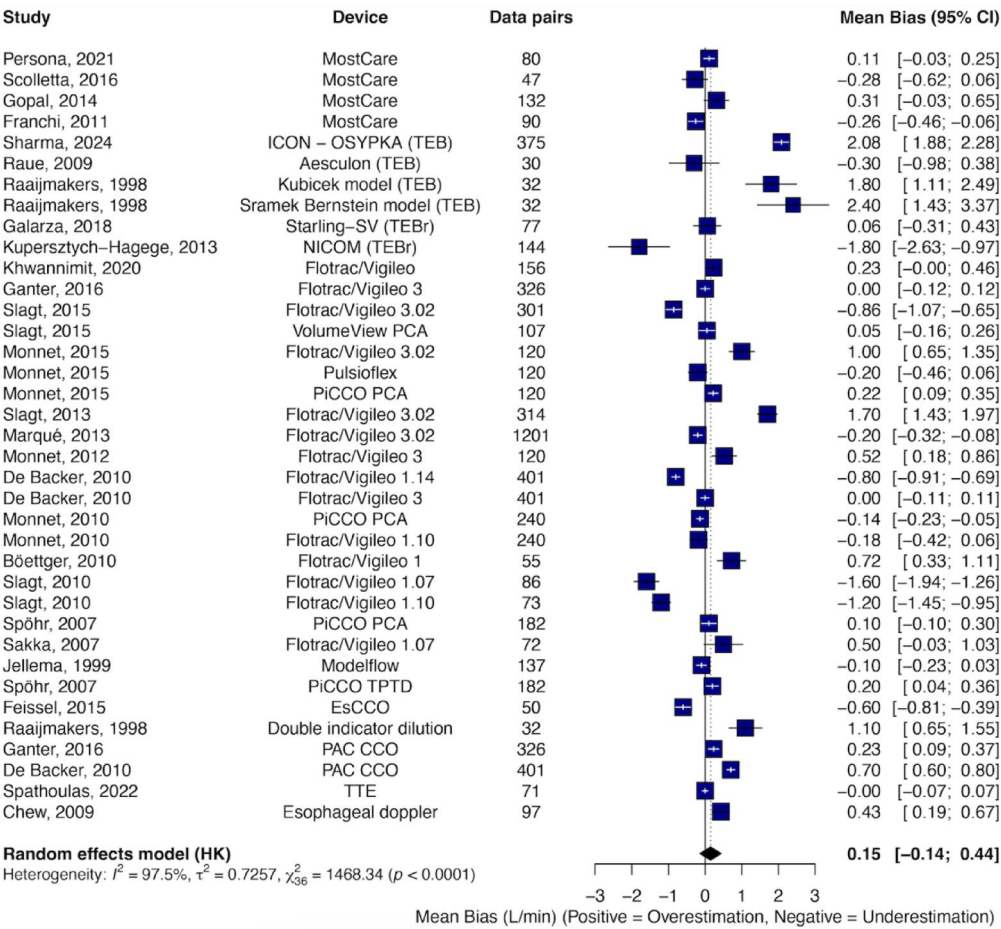


Fig. 2 Forest plot displaying the mean bias and 95% confidence intervals (CI) from studies comparing cardiac output (CO) measurements obtained using various CO monitoring devices versus reference methods. Data were derived from Bland–Altman analysis of the included studies. Data sorted by year and device type

bias of -0.03 L min^{-1} , LoA of $\pm 3.52 \text{ L min}^{-1}$, and a PE of 52.6% ($I^2 = 97.7\%$) (Fig. 4).

When stratified by reference method, pooled bias and PE were 0.23 L min^{-1} and 45.8% for PAC ($I^2 = 98\%$), 0.25 L min^{-1} and 55.7% for TPTD ($I^2 = 91.3\%$), and 0.40 L min^{-1} and 53.6% for TTE ($I^2 = 99.5\%$). A sensitivity analysis excluding four mixed-population studies did not alter pooled estimates (ESM-7) [24, 25, 32, 33].

Meta-regression did not identify any significant associations between mean bias and study-level variables, including mean CO, SVR, publication year, funding source, country, blinding, device type, or reference method (all $P > 0.46$). Residual heterogeneity remained high ($I^2 = 98.92\%$), suggesting the presence of unmeasured factors contributing to bias estimates.

Discussion

This systematic review and meta-analysis provides the most comprehensive synthesis to date of CO monitoring devices and techniques in ICU patients with septic shock. While some devices are no longer commercially

available, their inclusion offers important historical context, and highlights the persistent limitations in agreement and performance assessment that continue to affect even newer-generation devices [21, 55]. Despite a pooled mean bias near zero (0.15 L min^{-1}), the pooled PE of 49.4% far exceeds the 30% threshold for acceptable agreement [29]. Only calibrated PCA devices met this threshold, with a pooled PE of 25%. Trending performance was even more limited—of the 15 datasets reporting trending, just three met the $\geq 90\%$ concordance threshold [37, 40, 43].

Validation of CO monitors continues to rely heavily on Bland–Altman analysis, which assesses mean bias and inter-patient variability of the bias (via LoA and PE). While foundational, this approach does not fully capture other key performance domains that are also critical for bedside decision-making in the ICU—specifically, precision, time response, and trending ability [22, 28, 31].

The 30% PE threshold proposed by Critchley and Critchley’s was based on theoretical modelling: if both test and reference methods have 20% precision, the

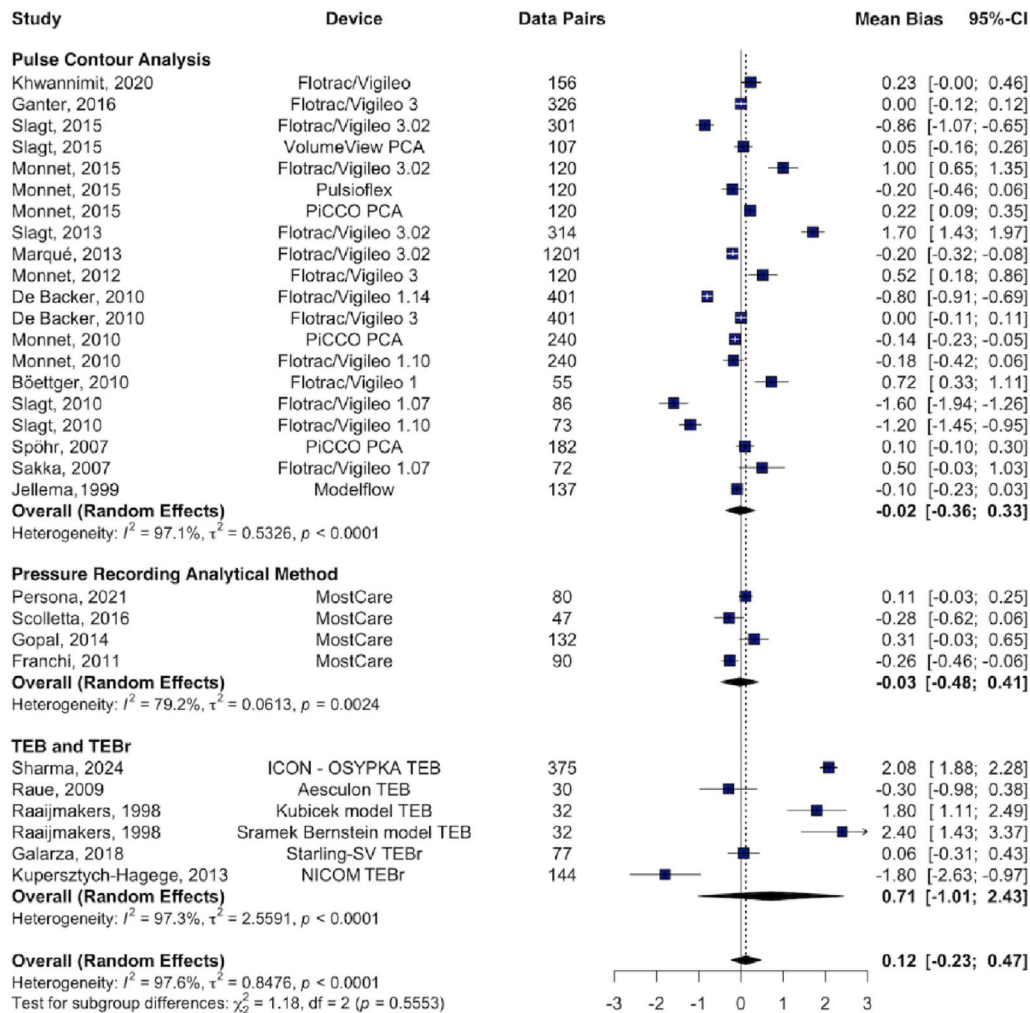


Fig. 3 Forest plot displaying the mean bias and 95% confidence intervals (CI) from studies comparing cardiac output (CO) measurements obtained using CO monitoring devices versus reference methods, stratified by device type. Data were derived from Bland–Altman analysis of the included studies. *PCA* Pulse Contour Analysis, *TEB* Thoracic Electrical Bioimpedance, *TEBr* Thoracic Electrical Bioreactance. Data sorted by year

expected PE is 28.3%, rounded to 30% for simplicity. This estimate is based on PAC-ITD precision measured under stable conditions [29, 56, 57]. Peyton et al. challenged this benchmark, citing studies that compared PAC-ITD to direct aortic transit-time flowmetry during periods of high CO variability—such as post-cardiopulmonary bypass and in pharmacologically destabilized pig models—which reported PE values of 42–49% [58, 59]. They argued for raising the agreement threshold to 45%. However, assessing agreement and precision of a snapshot method like PAC-ITD against a truly continuous reference requires synchronized measurement during stable phases to avoid distortion from physiological variability and response time mismatches [22]. Importantly, a PE threshold of 45% implies accepting devices that can only detect CO changes $\geq 40\%$ —a margin too wide to guide therapy in septic shock [60].

A major limitation of PE is that it conflates the imprecision of both the test and reference methods. Only three studies in this review quantified precision using the method proposed by Cecconi et al. [13, 42, 53]. Without this, PE becomes a blunt instrument. For instance, De Backer et al. reported a PE of 29.7% for FloTrac/Vigileo® version 3, but device precision was 28% based on a reference precision of 12%, rendering the PE estimate misleading [53].

Time response—defined as the delay between a hemodynamic intervention and the observed CO inflection—is arguably one of the most important yet neglected aspects of monitor performance. None of the included studies formally assessed the real-time responsiveness of these monitors during clinically relevant interventions such as passive leg raise or fluid challenges. Yet these dynamic hemodynamic challenges are among the most practical bedside applications of continuous CO

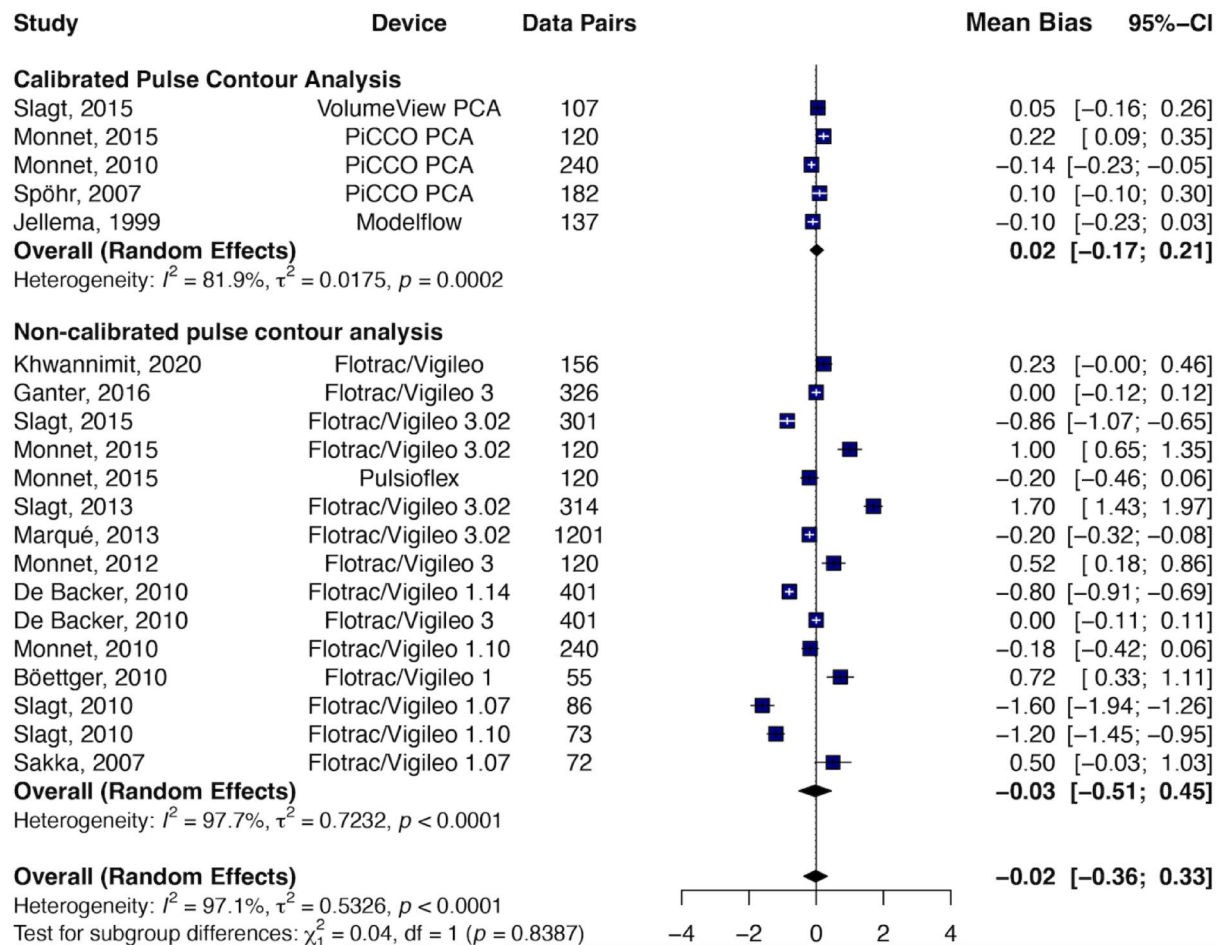


Fig. 4 Forest plot displaying the mean bias and 95% confidence intervals (CI) from studies comparing cardiac output (CO) measurements obtained using CO monitoring devices versus reference methods, stratified by pulse contour analysis device type (calibrated vs. non-calibrated). Data sorted by year

monitoring, especially in septic shock, where absolute CO values are less informative than timely detection of meaningful directional changes [61, 62]. Moreover, system specifications such as processing time and sampling frequency only account for part of the time response; actual responsiveness must be empirically tested. A monitor may process signals quickly but still delay displaying results due to internal filtering or averaging algorithms [22, 31]. Would a CO monitor that lags during a mini-fluid challenge be clinically useful? Likely not. Despite its centrality to bedside use, time response remains largely unstudied and should be formally incorporated into future validation frameworks.

Trending, defined as a device's ability to track directional CO changes, was also poorly assessed. Two of the three studies meeting the $\geq 90\%$ concordance threshold used smaller-than-recommended exclusion zones, potentially inflating trending performance by capturing physiologically irrelevant noise [37, 40]. Trending analysis with polar or 4-quadrant plots requires four key elements: (1) exclusion zones derived from recommended standards

[30] or device precision; [27, 29] (2) interventions that produce CO changes exceeding the exclusion zone; (3) sufficient data pairs post-exclusion to preserve statistical power; and (4) synchronization of test and reference response times. Most studies failed on all fronts [35, 40, 41, 43]. For instance, in evaluating PRAM, Persona et al. used small norepinephrine dose adjustments (0.03 mcg/kg/min) and reported poor trending, while Franchi et al.—using a MAP increase from 70 to 90 mmHg—reported 100% concordance, although the intervention was likely unrealistic for routine clinical practice [35, 37]. Furthermore, CO-guided interventions are typically recommended after initial resuscitation [6, 7], when patients may have limited CO reserve, further reducing Δ CO and increasing the risk of false-negative trending assessments. These methodological limitations raise concern about the strength of current confidence in CO monitor trending. Much of this confidence appears to rely on analyses unsuited to assess trending, such as ROC analysis [28, 63].

Ideally, reference methods should be continuous, automatic, and accurate. In reality, all are snapshot based, introducing variability in measurement timing and operator technique. In our subgroup analysis, echocardiography—despite being the most operator-dependent method—had a pooled PE comparable to TPTD. Evidence regarding its interchangeability with PAC-ITD is mixed, with some studies suggesting agreement [16] and others—particularly those involving transesophageal echocardiography—disputing it [64]. PAC-CCO, although labeled “continuous,” is in fact a semi-continuous averaging method with response times up to 10 min [65, 66]. In our review [41, 47, 52, 53] and prior meta-analysis [65], it consistently underperformed as both a test and reference method. Ultimately, PAC-CCO delay renders it unsuitable for evaluating rapid hemodynamic changes [22]. Given these inherent limitations—delayed signal responsiveness in PAC-CCO and substantial intra- and interobserver variability in TTE [18]—the conclusions drawn from individual validation studies and pooled agreement estimates using these references must be interpreted with exceptional caution. These methodological constraints likely contribute to variability across studies and underscore the need for more robust and physiologically responsive reference standards in future validation efforts.

From a clinical standpoint, device selection must balance performance with feasibility. Calibrated PCA systems such as PiCCO® and EV1000™ Edwards demonstrated the most consistent agreement but require both central arterial and venous access, along with regular recalibration in dynamic settings [15, 52, 67]. Notably, PiCCO® has shown comparable accuracy with radial or femoral artery cannulation, supporting broader applicability in the ICU [68]. Modelflow®, calibrated against PAC-ITD, showed good agreement [54], but is no longer commercially available, and its performance in unstable patients remains largely unstudied. Uncalibrated systems estimate CO relying on arterial waveform analysis and patient-specific demographic inputs to compute vascular tone; these assumptions can be disrupted in septic shock, particularly with low SVR, which has been shown to degrade agreement [69]. For example, FloTrac/Vigileo® version 3.02 underestimated CO unless SVR exceeded 700 dyn·s·cm⁻⁵, with PE improving from >50–32%. [49] Among uncalibrated systems, PRAM (MostCare™) appears promising; however, formal precision validation is lacking and trending performance remains inconsistent. As discussed, only Franchi et al. showed acceptable trending, but concordance was likely overestimated due to a small exclusion zone and an unusually large intervention effect size [37]. Moreover, PRAM is not plug-and-play—waveform quality must be continuously verified and optimized. TEB and TEBr consistently demonstrated

the poorest agreement (PE ≥ 60%) and should not be used in septic shock.

Overall, our findings, consistent with prior meta-analyses [21, 55, 64, 65, 70–72], clearly reinforce long-standing concerns that traditional validation approaches are ill-suited for dynamic, real-world critical care [22, 27].

Limitations

Strengths of this review include a rigorous methodology, strict inclusion of prospective ICU studies in septic shock, and comprehensive coverage of CO devices. Nonetheless, substantial mean bias heterogeneity persisted despite meta-regression, subgroup and sensitivity analyses. This likely reflects a combination of clinical and methodological variability—ranging from differences in disease severity, hemodynamic profiles, vasopressor use, and ventilator settings to study-specific design.

Notably, heterogeneity in pooled mean bias may be inherent and has been consistently reported in prior meta-analyses of CO measurement modalities [21, 55, 64, 65, 70–72]. Even under ideal conditions (e.g., synchronized time scales during minimal CO variability), interpatient variation in bias is expected due to the interaction between device/reference precision and differences in baseline CO. Nevertheless, pooled mean bias remains a clinically meaningful benchmark, particularly when interpreted by device type.

PE, which adjusts for baseline CO, offers a more standardized metric that would likely be more homogeneous. However, only two studies reported 95% confidence intervals for PE [41, 53], limiting our ability to compute standard errors or assess heterogeneity for this key parameter. Additional limitations include the use of reference methods that are not continuous—raising concerns about synchronization artifacts—as well as limited reporting of device precision and time response. Finally, this review did not assess whether CO monitoring devices alter hemodynamic management or improve clinical outcomes when used independently.

Future directions

To advance future validation and device development efforts, we encourage researchers to evaluate agreement within distinct hemodynamic phenotypes—such as low vs. high vascular tone—and to consistently report 95% confidence intervals for key agreement metrics, including the SD of bias, LoA, and PE. This approach will enable more meaningful benchmarking of device agreement performance and support the development of algorithms tailored to diverse hemodynamic states [31].

Overall, while agreement as a surrogate for accuracy remains important, future validation must also prioritize additional parameters for clinical utility—specifically device precision, time response, and trending ability.

Each domain should be independently assessed under realistic clinical conditions. (1) Accuracy and precision should be tested during periods of low CO variability. (2) Time response should be quantified and plotted as the delay to CO inflection post-hemodynamic intervention. (3) Trending should be evaluated using adequately powered polar plot analyses with clearly defined exclusion zones, ideally based on device precision or ROC analysis. Without this framework, validation risks becoming an academic exercise divorced from clinical relevance.

Conclusion

Despite decades of development, most CO monitors demonstrate poor agreement with reference methods in septic shock. More importantly, their bedside utility remains unknown, as current validation practices—anchored in Bland–Altman analysis—do not evaluate whether these devices can detect meaningful CO changes in real time under the dynamic conditions of critical care. Key performance domains—precision, time response, and trending ability—are infrequently and insufficiently assessed. Future validation must move beyond agreement alone and adopt standardized, clinically relevant metrics to determine the true usability of CO monitors in the ICU.

Abbreviations

BSA	Body surface area
CI	Cardiac index
CO	Cardiac output
CV	Coefficient of variation
ΔCO	Change in cardiac output
ESM	Electronic supplementary material
ICU	Intensive care unit
IQR	Interquartile range
LoA	Limits of agreement
LVOT	Left ventricular outflow tract velocity time integral
MAP	Mean arterial pressure
PAC	Pulmonary artery catheter
PAC-CCO	Pulmonary artery catheter continuous cardiac output
PAC-ITD	Pulmonary artery catheter intermittent thermodilution
PCA	Pulse contour analysis
PE	Percentage error
PRAM	Pressure recording analytical method
PWTT	Pulse wave transit time
QUADAS	Quality Assessment of Diagnostic Accuracy Studies-2
ROC	Receiver operating characteristic
SD	Standard deviation
SVR	Systemic vascular resistance
TEB	Thoracic electrical bioimpedance
TEBr	Thoracic electrical bioimpedance
TPTD	Transpulmonary thermodilution
TTE	Transthoracic echocardiography

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13054-025-05547-9>.

Supplementary Material 1.

Acknowledgements

Not applicable.

Author contributions

R.L. conceived the study, developed the protocol, supervised the systematic review process, verified data accuracy, conducted the statistical analysis, drafted the main manuscript, tables and figures. L.O. contributed to study design, data screening, extraction, and risk of bias assessment, and assisted in manuscript drafting. K.D.H. assisted with data screening, extraction, and risk of bias assessment. S.A.G. and G.U.M. contributed to data screening and extraction. M.C.S. served as the information specialist, developed the search strategy, and contributed to the methodology section. J.M. provided expert interpretation of device performance data. M.C. advised on study design and critically reviewed the manuscript. A.D. supervised the project, contributed to the conceptual framework, and revised the manuscript for important intellectual content. M.T.S. co-supervised the study, guided statistical methodology, and critically revised the manuscript. All authors reviewed and approved the final version of the manuscript.

Funding

None.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

This study is a systematic review and meta-analysis of previously published literature and did not require approval from an institutional review board or consent to participate.

Consent for publication

Not applicable. This study does not contain any individual person's data in any form (including individual details, images, or videos).

Competing interests

The authors declare no competing interests.

Received: 9 June 2025 / Accepted: 6 July 2025

Published online: 12 July 2025

References

- De Backer D, Cecconi M, Chew MS, et al. A plea for personalization of the hemodynamic management of septic shock. *Crit Care*. 2022;26(1):1–13. <https://doi.org/10.1186/s13054-022-04255-y>.
- Pinsky MR. Hemodynamic evaluation and monitoring in the ICU. *Chest*. 2007;132(6):2020–9. <https://doi.org/10.1378/chest.07-0073>.
- De Backer D, Aissaoui N, Cecconi M, et al. How can assessing hemodynamics help to assess volume status? *Intensive Care Med*. 2022;48(10):1482–94. <https://doi.org/10.1007/s00134-022-06808-9>.
- Douglas IS, Alapat PM, Corl KA, et al. Fluid response evaluation in Sepsis hypotension and shock: A randomized clinical trial. *Chest*. 2020;158(4):1431–45. <https://doi.org/10.1016/j.chest.2020.04.025>.
- Fleischmann-Struzek C, Mellhammar L, Rose N, et al. Incidence and mortality of hospital- and ICU-treated sepsis: results from an updated and expanded systematic review and meta-analysis. *Intensive Care Med*. 2020;46(8):1552–62. <https://doi.org/10.1007/s00134-020-06151-x>.
- Evans L, Rhodes A, Alhazzani W, et al. Surviving Sepsis campaign: international guidelines for management of Sepsis and septic shock 2021. *Crit Care Med*. 2021;49(11):e1063–143. <https://doi.org/10.1097/CCM.0000000000000537>.
- Cecconi M, De Backer D, Antonelli M, et al. Consensus on circulatory shock and hemodynamic monitoring. Task force of the European society of intensive care medicine. *Intensive Care Med*. 2014;40(12):1795–815. <https://doi.org/10.1007/s00134-014-3525-z>.
- Narang N, Thibodeau JT, Parker WF, et al. Comparison of accuracy of Estimation of cardiac output by thermodilution versus the Fick method using

- measured oxygen uptake. *Am J Cardiol.* 2022;176:58–65. <https://doi.org/10.1016/j.amjcard.2022.04.026>.
9. Schaffartzik W, Sanft C, Schaefer JH, Spies C. Different dosages of Dobutamine in septic shock patients: determining oxygen consumption with a metabolic monitor integrated in a ventilator. *Intensive Care Med.* 2000;26(12):1740–6. <https://doi.org/10.1007/s001340000635>.
 10. De Backer D, Vincent JL. The pulmonary artery catheter: is it still alive? *Curr Opin Crit Care.* 2018;24(3):204. <https://doi.org/10.1097/MCC.0000000000000502>.
 11. Reuter DA, Huang C, Edrich T, Shernan SK, Eltzschig HK. Cardiac Output Monitoring Using Indicator-Dilution...: Anesthesia & Analgesia. Accessed May 18, 2025. https://journals.lww.com/anesthesia-analgesia/fulltext/2010/03000/cardiac_output_monitoring_using_indicator_dilution.31.aspx
 12. Monnet X, Teboul JL. Transpulmonary thermodilution: advantages and limits. *Crit Care.* 2017;21(1):1–12. <https://doi.org/10.1186/s13054-017-1739-5>.
 13. Sakka SG, Kozieras J, Thuemer O, van Hout N. Measurement of cardiac output: a comparison between transpulmonary thermodilution and uncalibrated pulse contour analysis. *Br J Anaesth.* 2007;99(3):337–42. <https://doi.org/10.1093/bja/aem177>.
 14. Matthieu B, Karine NG, Vincent C, et al. Cardiac output measurement in patients undergoing liver transplantation: pulmonary artery catheter versus uncalibrated arterial pressure waveform analysis. *Anesth Analg.* 2008;106(5):1480. <https://doi.org/10.1213/ane.0b013e318168b309>.
 15. Monnet X, Anguel N, Naudin B, Jabot J, Richard C, Teboul JL. Arterial pressure-based cardiac output in septic patients: different accuracy of pulse contour and uncalibrated pressure waveform devices. *Crit Care Lond Engl.* 2010;14(3):R109. <https://doi.org/10.1186/cc9058>.
 16. Mercado P, Maizel J, Beyls C, et al. Transthoracic echocardiography: an accurate and precise method for estimating cardiac output in the critically ill patient. *Crit Care Lond Engl.* 2017;21(1):136. <https://doi.org/10.1186/s13054-017-1737-7>.
 17. Monnet X, Lai C, Backer DD. Why do we use transpulmonary thermodilution and pulmonary artery catheter in severe shock patients? *Ann Intensive Care.* 2025;15(1). <https://doi.org/10.1186/s13613-024-01400-4>.
 18. Jozwiak M, Mercado P, Teboul JL, et al. What is the lowest change in cardiac output that transthoracic echocardiography can detect? *Crit Care.* 2019;23(1):1–10. <https://doi.org/10.1186/s13054-019-2413-x>.
 19. Porter TR, Shillcutt SK, Adams MS, et al. Guidelines for the use of echocardiography as a monitor for therapeutic intervention in adults: a report from the American society of echocardiography. *J Am Soc Echocardiogr Off Publ Am Soc Echocardiogr.* 2015;28(1):40–56. <https://doi.org/10.1016/j.echo.2014.09.09>.
 20. Teboul JL, Saugel B, Cecconi M, et al. Less invasive hemodynamic monitoring in critically ill patients. *Intensive Care Med.* 2016;42(9):1350–9. <https://doi.org/10.1007/s00134-016-4375-7>.
 21. Joosten A, Desebbe O, Suehiro K, et al. Accuracy and precision of non-invasive cardiac output monitoring devices in perioperative medicine: a systematic review and meta-analysis. *Br J Anaesth.* 2017;118(3):298–310. <https://doi.org/10.1093/bja/aew461>.
 22. Squara P, Cecconi M, Rhodes A, Singer M, Chiche JD. Tracking changes in cardiac output: methodological considerations for the validation of monitoring devices. *Intensive Care Med.* 2009;35(10):1801–8. <https://doi.org/10.1007/s00134-009-1570-9>.
 23. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and Meta-Analyses of studies that evaluate health care interventions: explanation and elaboration. *PLOS Med.* 2009;6(7):e1000100. <https://doi.org/10.1371/journal.pmed.1000100>.
 24. Monnet X, Anguel N, Jozwiak M, Richard C, Teboul JL. Third-generation flotract/vigileo does not reliably track changes in cardiac output induced by norepinephrine in critically ill patients. *Br J Anaesth.* 2012;108(4):615–22. <https://doi.org/10.1093/bja/aer491>.
 25. Kupersztich-Hagege E, Teboul JL, Artigas A, et al. Bioreactance is not reliable for estimating cardiac output and the effects of passive leg Raising in critically ill patients. *Br J Anaesth.* 2013;111(6):961–6. <https://doi.org/10.1093/bja/aet282>.
 26. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529–36. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>.
 27. Cecconi M, Rhodes A, Poloniecki J, Della Rocca G, Grounds RM. Bench-to-bedside review: the importance of the precision of the reference technique in method comparison studies—with specific reference to the measurement of cardiac output. *Crit Care Lond Engl.* 2009;13(1):201. <https://doi.org/10.1186/cc7129>.
 28. Critchley LA, Lee A, Ho AMH. A critical review of the ability of continuous cardiac output monitors to measure trends in cardiac output. *Anesth Analg.* 2010;111(5):1180–92. <https://doi.org/10.1213/ANE.0b013e3181f08a5b>.
 29. Critchley LA, Critchley JA. A meta-analysis of studies using bias and precision statistics to compare cardiac output measurement techniques. *J Clin Monit Comput.* 1999;15(2):85–91. <https://doi.org/10.1023/a:1009982611386>.
 30. Critchley LA, Yang XX, Lee A. Assessment of trending ability of cardiac output monitors by Polar plot methodology. *J Cardiothorac Vasc Anesth.* 2011;25(3):536–46. <https://doi.org/10.1053/jjvca.2011.01.003>.
 31. Montenij LJ, Buhre WF, Jansen JR, Kruitwagen CL, de Waal EE. Methodology of method comparison studies evaluating the validity of cardiac output monitors: a Stepwise approach and checklist. *Br J Anaesth.* 2016;116(6):750–8. <https://doi.org/10.1093/bja/aew094>.
 32. Scolletta S, Franchi F, Romagnoli S, et al. Comparison between Doppler-Echocardiography and uncalibrated pulse contour method for cardiac output measurement: A multicenter observational study. *Crit Care Med.* 2011;44(7):1370–9. <https://doi.org/10.1097/CCM.00000000000001663>.
 33. Galarza L, Mercado P, Teboul JL, et al. Estimating the rapid haemodynamic effects of passive leg Raising in critically ill patients using bioreactance. *Br J Anaesth.* 2018;121(3):567–73. <https://doi.org/10.1016/j.bja.2018.03.013>.
 34. Yang XX, Critchley LA, Joynt GM. Determination of the precision error of the pulmonary artery thermodilution catheter using an in vitro continuous flow test rig. *Anesth Analg.* 2011;112(1):70. <https://doi.org/10.1213/ANE.0b013e3181ff475e>.
 35. Persona P, Valeri I, Saraceni E, De Cassai A, Calabrese F, Navalesi P. Cardiac output evaluation on septic shock patients: comparison between calibrated and uncalibrated devices during vasopressor therapy. *J Clin Med.* 2021;10(2):213. <https://doi.org/10.3390/jcm10020213>.
 36. Gopal S, Do T, Pooni JS, Martinelli G. Validation of cardiac output studies from the mostcare compared to a pulmonary artery catheter in septic patients. *Minerva Anestesiol.* 2014;80(3):314–23.
 37. Franchi F, Silvestri R, Cubattoli L, et al. Comparison between an uncalibrated pulse contour method and thermodilution technique for cardiac output Estimation in septic patients. *Br J Anaesth.* 2011;107(2):202–8. <https://doi.org/10.1093/bja/aer123>.
 38. Raue W, Swierzy M, Koplin G, Schwenk W. Comparison of electrical velocimetry and transthoracic thermodilution technique for cardiac output assessment in critically ill patients. *Eur J Anaesthesiol.* 2009;26(12):1067–71. <https://doi.org/10.1097/EJA.0b013e32832bdf94>.
 39. Raaijmakers E, Faes TJ, Kunst PW, et al. The influence of extravascular lung water on cardiac output measurements using thoracic impedance cardiography. *Physiol Meas.* 1998;19(4):491–9. <https://doi.org/10.1088/0967-3334/19/4/004>.
 40. Khwannimit B, Jomsuriya R. Comparison the accuracy and trending ability of cardiac index measured by the fourth-generation of flotract with the PICCO device in septic shock patients. *Turk J Med Sci.* 2020;50(4):860–9. <https://doi.org/10.3906/sag-1909-58>.
 41. Ganter MT, Alhashemi JA, Al-Shabasy AM, et al. Continuous cardiac output measurement by un-calibrated pulse wave analysis and pulmonary artery catheter in patients with septic shock. *J Clin Monit Comput.* 2016;30(1):13–22. <https://doi.org/10.1007/s10877-015-9672-0>.
 42. Slagt C, Helmi M, Malagon I, Groeneveld ABJ. Calibrated versus uncalibrated arterial pressure waveform analysis in monitoring cardiac output with transpulmonary thermodilution in patients with severe sepsis and septic shock: an observational study. *Eur J Anaesthesiol.* 2015;32(1):5–12. <https://doi.org/10.1097/EJA.0000000000000173>.
 43. Monnet X, Vaquer S, Anguel N, et al. Comparison of pulse contour analysis by pulsiotex and vigileo to measure and track changes of cardiac output in critically ill patients. *Br J Anaesth.* 2015;114(2):235–43. <https://doi.org/10.1093/bja/aeu375>.
 44. Slagt C, Beute J, Hoeksema M, Malagon I, Mulder JWR, Groeneveld JAB. Cardiac output derived from arterial pressure waveform analysis without calibration vs. thermodilution in septic shock: evolving accuracy of software versions. *Eur J Anaesthesiol.* 2010;27(6):550–4. <https://doi.org/10.1097/EJA.0b013e3283333a92>.
 45. Feissel M, Aho LS, Georgiev S, et al. Pulse wave transit time measurements of cardiac output in septic shock patients: A comparison of the estimated continuous cardiac output system with transthoracic echocardiography. *PLoS ONE.* 2015;10(6):e0130489. <https://doi.org/10.1371/journal.pone.0130489>.

46. Chew HC, Devanand A, Phua GC, Loo CM. Oesophageal doppler ultrasound in the assessment of haemodynamic status of patients admitted to the medical intensive care unit with septic shock. *Ann Acad Med Singap.* 2009;38(8):699–703.
47. Marqué S, Gros A, Chimot L, et al. Cardiac output monitoring in septic shock: evaluation of the third-generation Flotrac-Vigileo. *J Clin Monit Comput.* 2013;27(3):273–9. <https://doi.org/10.1007/s10877-013-9431-z>.
48. Sharma S, Ramachandran R, Rewari V, Tripathi A. Evaluation of electrical cardiometry to assess fluid responsiveness in patients with acute circulatory failure: A comparative study with transthoracic echocardiography. *Indian J Crit Care Med Peer-Rev Off Publ Indian Soc Crit Care Med.* 2024;28(7):650–6. <https://doi.org/10.5005/jp-journals-10071-24753>.
49. Slagt C, de Leeuw MA, Beute J, et al. Cardiac output measured by uncalibrated arterial pressure waveform analysis by recently released software version 3.02 versus thermodilution in septic shock. *J Clin Monit Comput.* 2013;27(2):171–7. <https://doi.org/10.1007/s10877-012-9410-9>.
50. Böttger SF, Pavlovic D, Gründling M, et al. Comparison of arterial pressure cardiac output monitoring with transpulmonary thermodilution in septic patients. *Med Sci Monit Int Med J Exp Clin Res.* 2010;16(3):PR1–7.
51. Spathoulas K, Tsolaki V, Zakynthinos GE, et al. The role of left ventricular ejection fraction and left ventricular outflow tract Velocity-Time integral in assessing cardiovascular impairment in septic shock. *J Pers Med.* 2022;12(11):1786. <https://doi.org/10.3390/jpm12111786>.
52. Spöhr F, Hettrich P, Bauer H, Haas U, Martin E, Böttiger BW. Comparison of two methods for enhanced continuous circulatory monitoring in patients with septic shock. *Intensive Care Med.* 2007;33(10):1805–10. <https://doi.org/10.1007/s00134-007-0703-2>.
53. De Backer D, Marx G, Tan A, et al. Arterial pressure-based cardiac output monitoring: a multicenter validation of the third-generation software in septic patients. *Intensive Care Med.* 2011;37(2):233–40. <https://doi.org/10.1007/s00134-010-2098-8>.
54. Jellema WT, Wesseling KH, Groeneveld JAB, et al. Continuous cardiac output in septic shock by simulating a model of the aortic input impedance: A comparison with bolus injection thermodilution. *Anesthesiology.* 1999;90(5):1317–28. <https://doi.org/10.1097/00005542-199905000-00016>.
55. Peyton PJ, Chong SW. Minimally invasive measurement of cardiac output during surgery and critical care: a meta-analysis of accuracy and precision. *Anesthesiology.* 2010;113(5):1220–35. <https://doi.org/10.1097/ALN.0b013e3181ee3130>.
56. Stetz CW, Miller RG, Kelly GE, Raffin TA. Reliability of the thermodilution method in the determination of cardiac output in clinical practice. *Am Rev Respir Dis.* 1982;126(6):1001–4. <https://doi.org/10.1164/arrd.1982.126.6.1001>.
57. Mackenzie JD, Haites NE, Rawles JM. Method of assessing the reproducibility of blood flow measurement: factors influencing the performance of thermodilution cardiac output computers. *Br Heart J.* 1986;55(1):14–24.
58. Bajorat J, Hofmockel R, Vagts DA, et al. Comparison of invasive and less-invasive techniques of cardiac output measurement under different haemodynamic conditions in a pig model. *Eur J Anaesthesiol.* 2006;23(1):23–30. <https://doi.org/10.1017/S0265021505001717>.
59. Botero M, Kirby D, Lobato EB, Staples ED, Gravenstein N. Measurement of cardiac output before and after cardiopulmonary bypass: comparison among aortic transit-time ultrasound, thermodilution, and noninvasive partial CO₂ rebreathing. *J Cardiothorac Vasc Anesth.* 2004;18(5):563–72. <https://doi.org/10.1053/j.jvca.2004.07.005>.
60. Cecconi M, Dawson D, Casaretti R, Grounds RM, Rhodes A. A prospective study of the accuracy and precision of continuous cardiac output monitoring devices as compared to intermittent thermodilution. *Minerva Anesthesiol.* 2010;76(12):1010–7.
61. Cecconi M, Parsons AK, Rhodes A. What is a fluid challenge? *Curr Opin Crit Care.* 2011;17(3):290. <https://doi.org/10.1097/MCC.0b013e32834699cd>.
62. Vincent JL, Cecconi M, De Backer D. The fluid challenge. *Crit Care.* 2020;24(1):703. <https://doi.org/10.1186/s13054-020-03443-y>.
63. Ruste M, Jacquet-Lagrèze M, Fellahi JL. Advantages and limitations of noninvasive devices for cardiac output monitoring: a literature review. *Curr Opin Crit Care.* 2023;29(3):259–67. <https://doi.org/10.1097/MCC.0000000000001045>.
64. Wetterslev M, Møller-Sørensen H, Johansen RR, Perner A. Systematic review of cardiac output measurements by echocardiography vs. thermodilution: the techniques are not interchangeable. *Intensive Care Med.* 2016;42. <https://doi.org/10.1007/s00134-016-4258-y>.
65. Kouz K, Michard F, Bergholz A, et al. Agreement between continuous and intermittent pulmonary artery thermodilution for cardiac output measurement in perioperative and intensive care medicine: a systematic review and meta-analysis. *Crit Care.* 2021;25(1):125. <https://doi.org/10.1186/s13054-021-03523-7>.
66. Haller M, Zollner C, Briegel J, Forst H. Evaluation of a new continuous thermodilution cardiac output monitor in critically ill patients: A prospective criterion standard study. *Crit Care Med.* 1995;23(5):860.
67. Friessecke S, Heinrich A, Abel P, Felix SB. Comparison of pulmonary artery and aortic transpulmonary thermodilution for monitoring of cardiac output in patients with severe heart failure: validation of a novel method. *Crit Care Med.* 2009;37(1):119–23. <https://doi.org/10.1097/CCM.0b013e31819290d5>.
68. De Wilde RBP, Breukers RBGE, Van Den Berg PCM, Jansen JRC. Monitoring cardiac output using the femoral and radial arterial pressure waveform. *Anaesthesia.* 2006;61(8):743–6. <https://doi.org/10.1111/j.1365-2044.2006.04712.x>.
69. Metzelder S, Coburn M, Fries M, et al. Performance of cardiac output measurement derived from arterial pressure waveform analysis in patients requiring high-dose vasopressor therapy. *Br J Anaesth.* 2011;106(6):776–84. <https://doi.org/10.1093/bja/aer066>.
70. Zhang Y, Wang Y, Shi J, Hua Z, Xu J. Cardiac output measurements via echocardiography versus thermodilution: A systematic review and meta-analysis. *PLoS ONE.* 2019;14(10):e0222105. <https://doi.org/10.1371/journal.pone.0222105>.
71. Sanders M, Servaes S, Slagt C. Accuracy and precision of non-invasive cardiac output monitoring by electrical cardiometry: a systematic review and meta-analysis. *J Clin Monit Comput.* 2020;34(3):433–60. <https://doi.org/10.1007/s10877-019-00330-y>.
72. Saugel B, Hoppe P, Nicklas JY, et al. Continuous noninvasive pulse wave analysis using finger cuff technologies for arterial blood pressure and cardiac output monitoring in perioperative and intensive care medicine: a systematic review and meta-analysis. *Br J Anaesth.* 2020;125(1):25–37. <https://doi.org/10.1016/j.bja.2020.03.013>.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.